Computer Science Department

TECHNICAL REPORT

VERY FAST ALGORITHMS FOR THE AREA OF THE UNION OF MANY CIRCLES

BY

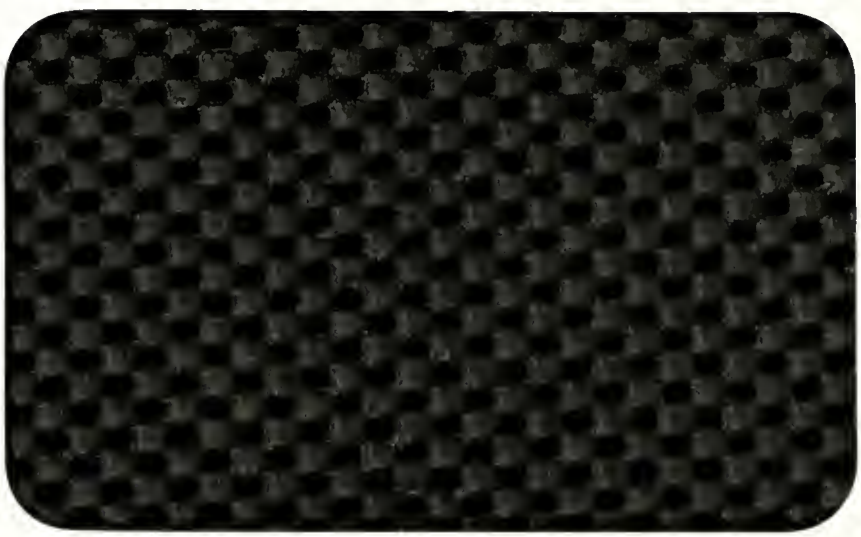
Paul G. Spirakis*

December 1983

Report #98

NEW YORK UNIVERSITY

Department of Computer Science
Courant Institute of Mathematical Sciences
251 MERCER STREET, NEW YORK, N.Y. 10012



VERY FAST ALGORITHMS FOR THE AREA OF THE
UNION OF MANY CIRCLES

BY

Paul G. Spirakis*

December 1983

Report #98

*This work was supported in part by the National Science Foundation Grant NSF-MCS83-00630.

C

ABSTRACT

We present here an $O(n^2)$ deterministic and an $O(n)$ probabilistic algorithm for computing the area of the union of n given circles. The accuracy of the probabilistic algorithm is controlled by the algorithm implementer (it can be set up to any desirable value). The probabilistic algorithm is optimal in the sense that $O(n)$ time is needed to compute the sum of n real numbers, under quite general models of computation. The probabilistic linear algorithm can be extended, in a straightforward manner, to compute the volume of the union of many similar objects (e.g. spheres) in k dimensions. Its time complexity is then $O(nk)$.

Our proof of fast convergence of the probabilistic method benefited a lot from a recent work of R. Karp and M. Luby on reliability. Our technique was also motivated by Karp and Luby's work, and some of our notation is heavily affected by their notation.

1. Introduction

The problem of estimating the area of the union of many circles in the plane was first posed by [Shamos, 78]. We present here an $O(n^2)$ deterministic and an $O(n)$ probabilistic algorithm for solving the problem. The recent work of [Karp, Luby, 83] on estimating the failure probability of an n-component system helped us a lot in stating the algorithm and formulating the proof of fast convergence. The probabilistic algorithm can be modified to compute the area of the union of other planar objects (of fixed description) e.g. triangles or boxes. The time is again $O(n)$. The algorithm can also be extended to compute the volume of the union of n spheres in k dimensions, in time $O(nk)$. The algorithm has similar time complexity for k-dimensional objects other than spheres, provided that each object has a fixed description (not dependent on n).

A very recent work of [Sharir, 83] can be extended to provide an $O(n \log^2 n)$ deterministic algorithm for computing the area of the union of n circles in the plane. The idea does not generalize easily to 3 dimensions. No efficient algorithms for the problem of computing volumes of unions of objects in more than 2 dimensions were presented in the past.

2. The problem in 2 dimensions

2.1 General remarks

The input is a list of n triples (x_i, y_i, r_i), $i = 1,\ldots,n$, where (x_i, y_i) are the Cartesian coordinates of the center of the circle C_i,

of radius r_i. The problem is to compute the area of the union $\bigcup_{i=1}^{n} C_i$. Let U be the area of the union. Let $E(C_i)$ denote the area of the circle C_i. Let $\sum = \sum_{i=1}^{n} E(C_i)$. The part of the plane covered by the union of the n circles is partitioned into nonoverlapping segments, whose boundaries are circular arcs (pieces of the circumferences of the circles C_i). Let $s_1, \ldots, s_m$ be an enumeration of these segments. Let $E(s_i)$ represent the area of segment s_i. Clearly, $U = \sum_{i=1}^{m} E(s_i)$. In our algorithms, we assume that all calculations are done in exact arithmetic.

Remark 2.1 The number of segments, m, is $O(n^2)$

Proof: Let us assume we have already i circles drawn and we next draw the $i+1^{th}$ circle. In the worst case, 2i new points of intersection of circumferences are introduced (2 for each of the i circles). Let us call them vertices. So, the total number of vertices for n circles is $v \leq \sum_{i=1}^{n} 2i = O(n^2)$. If we take out degeneracies, a vertex is shared by at most 4 segments, hence the total number of segments is $m \leq 4v = O(n^2)$. ●

The above remark leads to a straightforward $O(n^3)$ time and $O(n^3)$ space algorithm for computing U. This algorithm keeps a list of the segments created by the i first circles. Each segment is kept as a sequence of circular arcs, this sequence being the arcs which surround the segment, listed in the counterclockwise direction.
Each circular arc is a triple of a start point, an end point and a

curvature. Clearly, a segment created by i circles can have at most i surrounding arcs.

2.2 An $O(n^2)$ algorithm

This algorithm keeps, in each stage, just a list of circular arcs, surrounding the current union (of the first i cycles). The algorithm is as follows:

Algorithm 1

Initially: $i = 1$. Let $\ell(1)$ be the circumference of C_1. Do the following stage for $i = 2$ to n.

/* Begin stage i

Let $\ell(i-1)$ be the current list of arcs of the union (The arcs of each connected component are listed in counterclockwise direction). Find its intersections with the i^{th} circle (There can be $O(i)$ such intersections). Updatae the list by removing arcs which are now interior to the new union and by adding the surroundings of the part of the $i+1^{th}$ circle which is not covered by the union of the $i-1$ previous circles. The new list is the list $\ell(i)$.

/* End stage i.

Finally, compute the area enclosed into the arcs of the list $\ell(n)$.

<u>Lemma</u> <u>2.1</u> Algorithm 1 takes $O(n^2)$ time and $O(n)$ space.

<u>Proof</u>: The Algorithm needs $O(i)$ time, to update the list $\ell(i-1)$ at state i, since $\ell(i-1)$ may have up to $i-1$ arcs, one per each circle.

So, the total time spent during the n stages is $O(n^2)$. The area calculation takes $O(n)$ time. (The area calculation can be done by computing $O(n)$ integrals, one for each arc, and adding them with a plus or minus sign). The space needed is that of storing the list of arcs which surround the union. This list cannot have more than n arcs. ●

Note. Algorithm 1 computes not only the area of the union but also the boundary of the union of the n circles.

2.3 A straightforward Monte-Carlo approach.

2.3.1 Definitions - Terminology

Following [Karp, Luby, 83], let $\hat{U}$ be the estimate of the area of the union (U), produced by a randomized (Monte Carlo) algorithm MC.

Definition. MC is called an (ε,δ) algorithm of time $f(n)$, if $\forall \varepsilon,\delta \in (0,1)$, one can select a running time $f(n)$ (n = number of circles) of MC, such that the estimate $\hat{U}$ produced has the property

$$\text{Prob}\{|\frac{\hat{U}-U}{U}| > \varepsilon\} < \delta$$

(See [Karp, Luby 83] for a similar definition) If $f(n) = O(n)$, MC is a linear (ε,δ) algorithm.

Definition. Given the set of n circles C_i, $i = 1,\ldots,n$ in the plane, and a point P, let cover(P) be the number of circles to which P belongs (we consider circles to be closed sets).

2.3.2. The straightforward approach

Let us assume that we can find (in time $a(n)$) a closed curve A (enclosing area $E(A)$), which completely encloses the n circles. Assume also that we have a fast way (of time $\tau(n)$ per point selection) of selecting points within A (or the interior of A) in a uniformly random manner (meaning that the probability of a point selected to fall into an elementary area, $\Delta\varepsilon$, around a specific point P_0, is the same $\forall P_0$ (A plus interior (A)) and it is equal to $\frac{1}{E(A)} \cdot \Delta\varepsilon$.)

Finally, let $\beta(n)$ be the time it takes to decide if a given (arbitrarily selected) point P belongs to the union of the n circles.

Straightforward MC

(I) We select $N \geq \frac{1}{\varepsilon^2\delta}\left(\frac{E(A)}{U} - 1\right)$ points, uniformly randomly. (Note that, to do this selection we don't need to know U, just an upper bound on $E(A)/U$ is sufficient).

(II) For each point selected in (I), we test if it belongs to the union of circles. Let $M \leq N$ be the number of points found to belong in the union.

(III) Output $\hat{U} = \frac{M}{N} \cdot E(A)$.

Lemma 2.2. The straightforward MC method is an (ε,δ) algorithm of time $O\left(a(n) + \tau(n)\beta(n)\left(\frac{E(A)}{U} - 1\right)\frac{1}{\varepsilon^2\delta}\right)$.

Proof. Let $x_i = 1$ iff he i^{th} selected point belongs to the union, and $x_i = 0$ else, for $i = 1,\ldots,n$. Then $x_1+\ldots+x_N = M$ and $\mathrm{mean}(M) = \sum_{i=1}^{N}$ mean

$(x_i) = N \cdot \frac{U}{E(A)}$. So, mean $(\hat{U}) = \frac{E(A)}{N}$ mean$(M) = \frac{E(A)}{N} N \frac{U}{E(A)} = U(*)$; hence $\hat{U}$ is unbiased. Let σ^2 = variance (x_i) $i = 1,\ldots,N$. By elementary calculation, $\sigma^2 = \frac{U}{E(A)} - \frac{U^2}{E^2(A)}$ (**). Then, variance $(M) = \sigma^2 \cdot N$ hence

$$\text{variance}(\hat{U}) = \frac{\sigma^2 \cdot N}{N^2} E^2(A) = \frac{\sigma^2 E^2(A)}{N} = \frac{E(A) \cdot U - U^2}{N},$$

by (*), (**).

By Chebyshev Inequality then,

$$\text{Prob}\{|\hat{U}-U| > \varepsilon U\} < \frac{\text{variance}(\hat{U})}{\varepsilon^2 U^2} = \frac{E(A)U-U^2}{N\varepsilon^2 U^2} = \frac{1}{N\varepsilon^2}\left(\frac{E(A)}{U} - 1\right)$$

Hence $\text{Prob}\{\frac{|\hat{U}-U|}{U} > \varepsilon\} < \delta$, QED .

Clearly the running time is $a(n)$ (to select A) plus $O(N \cdot \tau(n) \cdot \beta(n))$. •

2.4 How to improve on the straighforward MC.

(An $O(n\log n)$ (ε,δ) algorithm).

A reasonable value of $a(n)$ is $O(n)$ (Eg. we can arbitrarily select a center and then find the most distant of the centers of the C_i's to that center, add to the distance found the maximum radius and draw a circle for A).

If we don't allow preprocessing, then $\beta(n) = O(n)$, leading to a time complexity of $O\left(n \cdot \tau(n) \cdot \left(\frac{E(A)}{U} - 1\right)\right)$. It should be clear that if

(1) A is easily describable so that $\tau(n) = O(1)$ and

(2) $\frac{E(A)}{U} = O(1)$, then we would get a linear (ε,δ) algorithm.

Requirement (2) implies that E(A) is close to U. One such A would be the boundary set of the union $\bigcup_{i=1}^{n} C_i$ itself, but then the task of selecting points from the interior of A uniformly randomly becomes a hard task, since a point belonging to $k > 1$ circles cannot be counted as a "whole point" when computing M.

A first modification to MC (Call it Extended MC, EMC) is the following:

Extended MC

(1) Let A be the boundary of the union of the n circles. (i.e $E(A) = U$)

(2) To select a random point in the interior of A we select one of the C_i at random (Call it C_j) and then we select one random point P_j of the circle selected.

(3) We compute $M = \sum_{j=1}^{N} (\text{cover } (P_j))^{-1}$

(4) We output $\hat{U} = \frac{M}{N}$

It is relatively easy to prove that if $N > \frac{1}{\varepsilon^2 \delta}$ then EMC is an (ε,δ) algorithm. Its complexity depends on the time F(n) to find cover (P) for each P. By using the Voronoi diagram construction as a preprocessing idea, we can compute cover (P) in time O(log n) for a given P. This will result in an $O(n \log n)(\varepsilon,\delta)$ EMC algorithm.

In the next section we will improve upon the time complexity of the O(n log n) EMC by using the following two probabilistic ideas:

(a) We shall assign greater probability of selection to points belonging to circles of bigger area.

(b) We will estimate covers through a very fast probabilistic experiment.

3. A linear (ε,δ) algorithm.

3.1 Linear Monte Carlo (LMC) Algorithm Initialization:

c is a small constant > 20

S:=0 , number.of.trials:=0;

TIME:=0

Compute Σ = the sum of the areas of the circles, i.e.

$$\Sigma = \sum_{i=1}^{n} E(C_i)$$

STEP 1 Randomly select cycle C_i with probability $p_i = \frac{E(C_i)}{\Sigma}$

STEP 2 Randomly select a point s in C_i

STEP 3 f:=0

Repeat until a C_j is selected such that $s \in C_j$

BEGIN

Randomly select C_j with probability 1/n

f:=f+1

Test if $s \in C_j$

TIME := TIME +1

END

Let $\hat{U} := \frac{f}{n} \cdot \Sigma$

STEP 4 number•of•trials:=number•of•trials + 1

$S := S + \hat{U}$

if TIME < $\frac{c \cdot n}{\delta\varepsilon^2}$ then goto STEP 1

STEP 5 Output $U^* := \dfrac{S}{\text{number}\cdot\text{of}\cdot\text{trials}}$

End of algorithm LMC.

By the description of the algorithm it is clear that it stops in $O(\frac{n}{\delta\varepsilon^2})$ steps.

We wish to prove that LMC is an (ε,δ) algorithm, i.e. that

$$\text{Prob}\{|U^*-U| > \varepsilon U\} < \delta \qquad \text{i.e.}$$

$$\text{Prob}\{|S - T\cdot U| > \varepsilon TU\} < \delta \qquad (\text{IN1})$$

where T is the number·of·trials

3.2 Notation and some useful remarks.

Definition. Let $t(T)$ be the sum $f_1 + f_2 + \ldots + f_T$, i.e. the sum of f's of the first T trials.

Definition. Let $U^*(T)$ be the value of U^* computed if the algorithm makes T trials.

Definition. Let $\tilde{U}_i$ be the estimate of the area of the union for the i^{th} trial.

Definition. Let $T(t)$ be the number of trials done in TIME = t.

Definition. Let $U^*(t)$ be the output of algorithm LMC given TIME = t.

Note that the details of STEP 2 are as follows: The probability that the point s will fall into an elementary area dA of C_i is the same for all elementary areas and is equal to $\frac{1}{\pi R_i^2}$ dA (R_i = radius of C_i).

Let f_i be a random variable indicating the value of f at the end of the i^{th} trial in LMC.

Lemma 3.1

(a) The mean value of f_i is mean $(f_i) = \frac{nU}{\Sigma}$

(b) The mean value of f_i^2 is mean $(f_i^2) \leq \frac{2n^2U}{\Sigma}$.

Proof

(a) Let $E(C_j,s)$ be the event that cycle C_j is selected in STEP 1 of the i^{th} trial and a point is selected in C_j in the small area dA (in C_j) around $s \in C_j$.

We have

$$\mathrm{Prof}\{E(C_j,s)\} = \frac{E(C_j)}{\Sigma} \cdot \frac{1}{\pi R_j^2}\, dA$$

Also

$$\mathrm{Prob}\{f_i=k/E(C_j,s)\} = (1-p(\text{success}))^{k-1} \cdot p(\text{success})$$

where $p(\text{success}) = \frac{\text{cover}(s)}{n}$. We conclude that $\text{mean}\{fi/E(C_j,s)\} = \frac{n}{\text{cover}(s)}$.

But

$$\text{mean}(f_i) = \sum_{\text{all}(C_j,s)} \text{mean }(fi/E(C_j,s)) \cdot \frac{E(C_j)}{\Sigma} \frac{1}{\pi R_j^2} dA$$

i.e.

$$\text{mean}(f_i) = \frac{n}{\Sigma} \int_{\substack{\text{over all} \\ \text{circles } C_j}} \frac{E(C_j)}{\text{cover}(s)} \frac{1}{E(C_j)} dA$$

$$\Rightarrow \text{mean}(f_i) = \frac{n}{\Sigma} \int_{\substack{\text{over} \\ \text{all circles } C_j}} \frac{dA}{\text{cover}(s)}$$

But $\displaystyle\int_{\substack{\text{over} \\ \text{all circles}}} \frac{dA}{\text{cover}(s)}$ = area of union = U. So, mean $(f_i) = \frac{nU}{\Sigma}$. QED

(b) $$\text{mean }(f_i^2/E(C_j,s)) = \sum_{k=0}^{\infty} k^2(1-p(\text{success}))^{k-1} \cdot p(\text{success})$$

$$\leq \frac{2}{p^2(\text{success})}$$

(by basic properties of geometric distributions).

So

$$\text{mean}(f_i^2/E(C_j,s)) \leq \frac{2n^2}{\text{cover}^2(s)}$$

Hence

$$\text{mean}(f_i^2) \leq \int_{\text{all }(C_j,s)} \frac{2n^2}{\text{cover}^2(s)} \frac{E(C_j)}{\Sigma} \cdot \frac{1}{E(C_j)} dA$$

$$< \frac{2n^2}{\Sigma} \int_{\text{all } C_j\text{'s}} \frac{dA}{\text{cover}(s)} \left(\text{because } \frac{1}{\text{cover}^2(s)} < \frac{1}{\text{cover}(s)}\right)$$

So

$$\text{mean }(f_i^2) < \frac{2n^2U}{\Sigma} \qquad \text{QED}$$

Corollary 3.1

Let $\tilde{U}_i$ be the random variable indicating the value of $\tilde{U}$ in the i^{th} round. Then mean $(\tilde{U}_i) = U$ and mean $(\tilde{U}_i^2) < 2\ \Sigma \cdot U$

Proof

$$\tilde{U}_i = \frac{\Sigma}{n} \cdot f_i \Rightarrow \text{mean }(\tilde{U}_i) = \frac{\Sigma}{n} \cdot \frac{nU}{\Sigma} = U.$$

Also

$$\tilde{U}_i^2 = \frac{\Sigma^2}{n^2} f_i^2 \Rightarrow \text{mean }(\tilde{U}_i) < \frac{\Sigma^2}{n^2} \cdot \frac{2n^2U}{\Sigma}$$

(by using Lemma 3.1b)

$$\Rightarrow \text{mean }(\tilde{U}_i^2) < 2\Sigma U. \qquad \text{QED}$$

3.3 The proof that LMC is an (ε,δ) algorithm.

<u>Theorem 3.1</u> For any $\varepsilon,\delta \in (0,1)$

$$\mathrm{Prob}\{|U^*(t) - U| \geq \varepsilon U\} < \delta \quad \text{where}$$

$$t= \frac{cn}{\delta\varepsilon^2}, \text{ c a constant.}$$

<u>Proof</u>. We use here the techniques of [Karp, Luby 1983].

From the program of LMC

$$f_1 + f_2 + \ldots + f_T = t(T) \quad (T = \text{total number of trials}) \quad \text{i.e.}$$

$$T \cdot \mathrm{mean}\ (f_i) = \mathrm{mean}\ (t(T))$$

So, if $t = \dfrac{c_1 n}{\delta\varepsilon^2}$ for some c_1 constant then we expect to do

$$T = \frac{c_1}{\delta\varepsilon^2}\frac{\Sigma}{U}\frac{n}{n} = \frac{c_1}{\delta\varepsilon^2}\frac{\Sigma}{U} \leq \frac{c_1 n}{\delta\varepsilon^2}$$

number of trials (because $\frac{\Sigma}{U} \leq n$ *)). Let k be equal to $\frac{c_1}{\delta\varepsilon^2}\frac{\Sigma}{U}$. Put $t' = \frac{c_1 n}{\delta\varepsilon^2}$ and $t'' = t'(1+\beta)$ where β is a constant, to be determined.

Let us estimate the

$$\mathrm{Prob}\left\{ \frac{|U^*(t'')-U|}{U} \geq \varepsilon \right\}$$

This is equal to

---

* $\Sigma \leq n \cdot \max_i E(C_i) \leq nU$

$$\text{Prob}\{|U^*(t'')-U| \geq \varepsilon U \text{ and } T(t'') < k\} +$$

$$\text{Prob}\{|U^*(t'') - U| \geq \varepsilon U \text{ and } T(t'') \geq k\} \qquad \text{(EQA)}$$

But

$$\text{Prob}\{|U^*(t'') - U| \geq \varepsilon U \text{ and } T(t'') < k\}$$

$$\leq \text{Prob}\{T(t'') < k\} \qquad \text{(EQB)}$$

But $T(t'') < k \iff f_1 + f_2 + \ldots + f_k > t''$. So

$$\text{Prob}\{T(t'') < k\} = \text{Prob}\{f_1 + \ldots + f_k > t''\}$$

$$= \text{Prob}\left\{\frac{f_1+\ldots+f_k}{k} > \frac{t'(1+\beta)}{k}\right\}$$

So

$$\text{Prob}\{T(t'') < k\} \leq \text{Prob}\left\{\left|\frac{\frac{f_1+\ldots+f_k}{k} - \frac{t'}{k}}{(t'|k)}\right| > \beta\right\}$$

$$\leq \frac{2\delta\varepsilon^2}{c_1\beta^2} \text{ (by Cheyshev inequality).}$$

(This is because $k = \frac{c_1}{\delta\varepsilon^2}\frac{\Sigma}{U}$ and mean $\left(\frac{f_1+\ldots+f_k}{k}\right)$ = mean$(f_i) = \frac{nU}{\Sigma}$ and $t' = \frac{c_1 n}{\delta\varepsilon^2}$.

Hence $\frac{t'}{k} = \frac{nU}{\Sigma} = \text{mean}\ \frac{f_1+\ldots+f_k}{k})\ )$.

From (EQB) we finally have

$$\text{Prob}\{|U^*(t'') - U| > \varepsilon U \text{ and } T(t'') < k\} < \frac{2\delta\varepsilon^2}{c_1\beta^2} \qquad \text{(EQC)}$$

Also

$$\text{Prob}\{|U^*(t'') - U| > \varepsilon U \text{ and } T(t'') > k\}$$

$$= \text{Prob}\{\exists\ r > k\text{: } T(t'') > k \text{ and } |\frac{\tilde{U}_1+\ldots+\tilde{U}_r}{r} - U| > \varepsilon\ U\}$$

$$< \text{Prob}\{\exists\ r > k\text{: } |\frac{\tilde{U}_1+\ldots+\tilde{u}_r}{r} - U| > \varepsilon U\} \qquad \text{(EQD)}$$

Here we use the following fact <u>Kolmogorov's</u> <u>inequality</u> (see [Feller,57])

Let $Y_1,\ldots,Y_m$ be independent random variables with the same distribution as Y and let mean(Y) and mean(Y^2) be finite. Let $\sigma^2(Y)$ be the variance of Y.

Let $R_\ell = \sum_{i=1}^{\ell} Y_i$, then

$$\text{Prob}\{\ \exists \ell\text{: } 1 < \ell < n \text{ and } |R_\ell - \ell\cdot\text{mean}(Y)| > \varepsilon\cdot m\cdot\text{mean}(Y)\}$$

$$< \frac{\sigma^2(Y)}{\varepsilon^2\cdot m\ \text{mean}^2(Y)}$$

By applying the above inequality for $Y_i = \tilde{U}_i$, $i=1,\ldots,T'$, $m = T'$, we get

$$\text{Prob}\left\{\ \ell : \left| \frac{\frac{\tilde U_i + \ldots + \tilde U_\ell}{\ell} - U}{U} \right| > \frac{\varepsilon}{\ell} \cdot T' \right\} < \frac{2\Sigma}{\varepsilon^2 T' U}$$

(To verify this, put mean $(\tilde U_i) = U$ $\forall_i$- and $Y = \tilde U_i$ (some i) and $\frac{\sigma^2(Y)}{U^2} <$ $\frac{\text{mean}(Y^2)}{U^2} < 2\,\frac{\Sigma}{U}$) (T′ is arbitrary). Let us now use $T' = 2^i \cdot T$ $i=1,\ldots,\infty$.
Then we get

$$P_1 = \text{Prob}\left\{\ \ell :\ 1 < \ell < N \cdot 2^i \text{ and } \left| \frac{\frac{\tilde U_i + \ldots + \tilde U_\ell}{\ell} - U}{U} \right| > 2\varepsilon \right\}$$

So we conclude

$$\text{Prob}\left\{\ \ell :\ N\ 2^{i-1} < \ell < N \cdot 2^i \text{ and } \left| \frac{\frac{\tilde U_1 + \ldots + \tilde U_\ell}{\ell} - U}{U} \right| > 2\varepsilon \right\}$$

$$< \frac{2\Sigma}{U} \frac{1}{\varepsilon^2 N 2^i} \qquad \text{(EQE)}$$

Now, we have

$$\text{Prob}\left\{ \left| \frac{\frac{\tilde U_1 + \ldots + \tilde U_\ell}{\ell} - U}{U} \right| < \varepsilon \text{ for some } \ell > k \right\}$$

$$< \sum_{i=1}^{\infty} \text{Prob}\left\{\ \ell :\ k \cdot 2^{i-1} < \ell < k \cdot 2^i \text{ and } \left| \frac{\frac{\tilde U_1 + \ldots + \tilde U_\ell}{\ell} - U}{U} \right| > 2 \cdot \frac{\varepsilon}{2} \right\}$$

$$< \sum_{i=1}^{\infty} \frac{2\Sigma}{U} \frac{1}{(\varepsilon/2)^2 k} \frac{1}{2^i} \quad \text{(by EQE)}$$

$$< \sum_{i=1}^{\infty} \frac{2\Sigma}{U} \frac{4}{\varepsilon^2} \frac{1}{2^i} \cdot \frac{\delta\varepsilon^2}{c_1} \frac{U}{\Sigma} \quad \text{(use value of k)}$$

$$< \frac{8\delta}{c_1}$$

That is, we just proved

$$\text{Prob}\left\{ \text{For some } \ell > k, \left| \frac{\tilde{U}_1+\ldots+\tilde{U}_\ell}{\ell} - U \right| > \varepsilon U \right\} < \frac{8\delta}{c_1} \qquad \text{(EQF)}$$

By (EQA), (EQB), (EQC), (EQD), (EQF) we now have

$$\text{Prob}\left\{\frac{U^*(t'')-U}{U} > \varepsilon\right\} < \frac{2\delta\varepsilon^2}{c_1\beta^2} + \frac{8\delta}{c_1}$$

Let $\delta' = \frac{\delta}{c_1}\left(\frac{2\varepsilon^2}{\beta^2} + 8\right)$. Then, if $c = c_1(1+\beta)$, we have proved that for any ε, $\delta' \in (0,1)$

$$\text{Prob}\{|U^*(t'') - U| > \varepsilon\, U\} < \delta'$$

where $t'' = \frac{cn}{\delta'\varepsilon^2}$, c a constant i.e. LMC is an (ε,δ') algorithm. QED

4. Applications to intersection problems

One can use the linear probabilistic estimation of union in order to check (in linear time) if the area (volume) of intersection of n cycles (spheres) is at least $\alpha \cdot \Sigma$ where $\alpha \in (0,1)$ and Σ is the sum of their areas (volumes).

Let U be the area (volume) of the union of the n objects. Clearly $U = \Sigma$ means the objects do not intersect (we are leaving out zero area intersections). We want to test efficiently if $\Sigma - U \geq \alpha\Sigma$ i.e. if $U \leq \Sigma(1-\alpha)$. We run the union estimation algorithm with a small ε (e.g. $\varepsilon = \frac{\alpha}{k}$, $k \geq 10$). If the estimated $U^* \leq \Sigma(1-\alpha) - \varepsilon$ then with probability at least $1-\delta$ the area of intersection is at least $\alpha \cdot \Sigma$.

The above technique can be extended to objects of higher dimensions.

References

[Feller, 57] An introduction to probability theory and its applications, Vol. I, Wiley 1957.

[Karp, Luby 1983] Monte Carlo Algorithms for Enumeration and Reliability Problems by R. Karp and M. Luby, 24th Annual Symposium on Foundations of Computer Science, Nov. 1983.

[Sharir, 83] M. Sharir, private communication.

NYU c.1
Comp. Sci. Dept.
TR-98 Spirakis
AUTHOR
Very fast algorithms for
TITLE
the area of the union of...

DATE DUE	BORROWER'S NAME